图书情报工作
第 62 卷 第 16 期 2018 年 8 月

我国图情领域社会化标签研究主要问题分析*

■ 李旭晖[1] 李媛媛[1] 马费成[1,2]
[1] 武汉大学信息管理学院 武汉 430072 [2] 武汉大学信息资源研究中心 武汉 430072

摘要：［**目的/意义**］社会化标签是我国图书情报与数字图书馆学科领域的研究热点，在近几年的发展过程中，对于标签的研究范围拓展至信息质量、信息检索、信息推荐等多个图情研究领域，对图情学科标签研究进展的分析或评述将为"社会化标签"未来的发展研究提供依据。［**方法/过程**］选取 2010 年至 2017 年间的图情领域期刊文献，利用 CiteSpace 可视化剖析社会化标签研究热点、发展进程、研究趋势，并从标签应用、标签信息、标签质量、信息推荐、信息检索 5 个角度对主题研究进行分析总结。［**结果/结论**］基于社会化标签研究现状及问题提出相关建议，如拓展标签在本体、信息检索层面的研究范围，以用户为中心适应个性化需求，将研究成果应用于实践等。

关键词：社会化标签 图书馆学 情报学
分类号：G250
DOI：10. 13266/j. issn. 0252 – 3116. 2018. 16. 014

1 引言

近年来，随着互联网技术的日益成熟，越来越多的社会资源被数据化，人们有了更多的渠道、方式获取信息，信息检索、分类等个性化需求与日俱增。最早将分类权限赋予用户的网络平台是标签类网站 Delicious，它允许用户对网站或网页内容添加标签，以实现网站资源的共享和自组织。Delicious 对资源分类权限的调整使得用户作为信息接收者的同时，也成为信息的生产者、描述者、组织者和共享者。相较于传统的信息组织方式，这种将用户标注作为信息资源分类的方式极大促进了信息的社会化，为信息个性化需求指明了新的发展方向。T. Vander Wal 于 2004 年提出 Folksonomy 的概念，意为大众分类法，并将其划分为广义自由分类法和狭义自由分类法两种，狭义与广义的区分在于标注者在知识结构、兴趣领域上同质与否。在此基础上，出现了"社会化标签""社会化标注""社会编目""社会元数据"等术语，这类标注术语的兴起体现了社会化标签的分类本质。

社会化标签是用户基于个人认知对信息资源添加的语义描述，作用于对 Web 资源的分类、组织和检索。社会化标签反映了客观信息和用户的主观认知，是用户实现资源组织、管理和共享的基础。在大众标注过程中，标签对信息资源进行了关键词分类、信息描述、注释、添加元数据等，体现了社会化标签所包含的客观信息与主观信息之间的交互融合。国内典型的社会化标签网站包括豆瓣网、微博、知乎等，作为用户原创内容（User Generated Content，UGC）产物，社会化标签融入了社会公众对信息的理解与认识，增强了用户与互联网信息之间的联系，为信息个性化服务、信息分类、信息推荐、信息行为等研究及应用带来了重要的数据资源。社会化标签在政府、企业、商业、教育、个人、其他公益类网站都有所应用，利用社会化标签，人们可以对网络资源进行收集标记、分类组织、分享推荐，通过标签浏览、检索、评价、添加、传播、共享等活动即可实现信息素养的全面发展[1]。

此外，标签还可用于信息检索、数据可视化、个性化服务、知识挖掘等应用领域，是应用领域极其广泛、研究范围拓展度极高的主题词。

* 本文系国家自然科学基金重大研究计划"大数据驱动的管理与决策研究"重点支持项目"基于知识关联的金融大数据价值分析、发现及协同创造机制"（项目编号：91646206）和教育部人文社会科学重点研究基地重大项目"大数据资源的语义表示与组织研究 – 面向文化遗产领域"（项目编号：16JJD870002）研究成果之一。
作者简介：李旭晖（ORCID：0000 – 0002 – 1155 – 3597），副教授，研究生导师，E-mail：lixuhui@whu. edu. cn；李媛媛（ORCID：0000 – 0001 – 7292 – 4366），硕士研究生；马费成，教授，博士生导师。
收稿日期：2018 – 03 – 12 修回日期：2018 – 05 – 21 本文起止页码：120 – 131 本文责任编辑：徐健

2 研究背景

2.1 研究问题的提出

由于生活背景、信息素养、学科知识等条件的不同,由用户创建的社会化标签具有形式随意性、主观性、不确定性等特点,标签特性使其在信息检索、分类、推荐等具体应用过程中产生质量不均、效率低下等问题。为了更好地利用社会化标签,适应用户获取和消费信息的需求,学术界对社会化标签的研究由对标签的实践应用、理论探讨拓展至与应用相关联的科学研究,包括可视化、标签表示、信息推荐、标签质量等。作为网络资源信息组织的新形式,社会化标签研究领域的拓展使得标签引发越来越多的关注,成为图书与情报学领域的研究热点。通过文献调查,图情领域相关学者对“社会化标签”的研究主要集中在信息检索、信息推荐、标签表示、知识网络等领域,鲜有对图情学科领域内的社会化标签研究情况的加以探讨。学科知识与研究框架的相互作用影响着专家学者对社会化标签相关问题的研究效果,掌控专业领域内的专题研究趋势或情况将有利于专家学者对社会化标签研究走向的掌控,推进信息科学的发展。笔者选取图情领域的“社会化标签”期刊文献,对研究热点、研究现状进行探讨分析,进而从更宏观层面了解该领域对“社会化标签”的研究情况。

2.2 研究对象的选择

本研究以“标签”“Folksonomy”“Social Tag”作为检索词,从中国知网、Sci-Hub 检索我国图书情报与数字图书馆学科领域期刊文献,年限设置为 2010 年至 2017 年,在剔除非相关文献后,共提取出 132 篇较具代表性的社会化标签主题文献。CiteSpace 通过文献关键词聚类以时间线视图的形式展示主题文献及各个子领域的研究演进过程,各节点半径体现关键词出现频次,节点间连线表示存在共现关系,节点按颜色色度变化展示不同年份,蓝色代表较早的年份,红色代表最近年份[2]。笔者选择此方法对“社会化标签”的期刊文献进行初步分析,关键词时间线视图如图 1 所示:

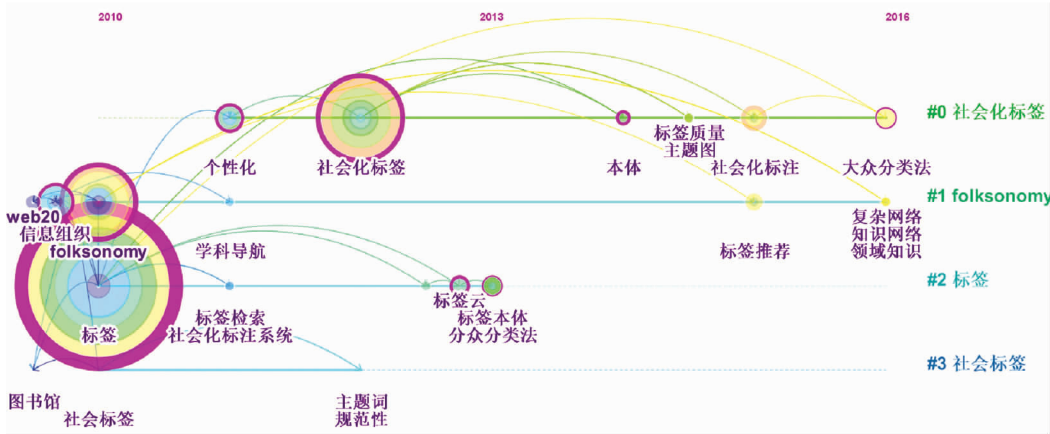


图 1 关键词时间线视图

由图 1 可知,在样本文献中,标签、Folksonomy、社会化标签是出现频次较高的关键词。最早的“社会化标签”的期刊文献发布于 2011 年左右,研究文献量在 2012 年达到高峰,研究内容包括个性化、本体、标签质量、主题图、社会化标注、大众分类法等;与“社会化标签”相关联的“Folksonomy”研究领域探讨了信息组织、学科导航、标签推荐、复杂网络、知识网络、领域知识等内容;相关的“标签”则探讨了标签检索、标签云、标签本体、分众分类法等;社会标签主要研究图书馆、主题词、规范性方面的内容。

2.3 研究热点分布

关键词共现网络分析也称作共词分析,通过分析文本内容信息获取共同出现的关键词,以探索学科领域中相关主题的关联关系。在关键词共现网络图中,圆圈越大代表关键词出现频次越多,关键词共现情况体现了关键词的中心性,中心性较高的关键词在共现网络中影响力越大,根据关键词出现频次及中心性情况可以了解文献主题的研究热点分布情况。通过相关文献调查分析,我国图情领域学者对社会化标签的分析研究领域跨度较广,涵盖自然科学、社会科学、人文科学等多个研究领域。根据图 2 的关键词共现网络,可知“Folksonomy”与“标签”关键词共现频次最高,样本文献中有 5 次。通过关键词标签聚类可知,标签与信息组织、信息检索、标注系统、本体等多个研究领域

关系紧密，并与 Folksonomy、社会标签等话题点构成社会化标签研究网络的主体部分。与社会化标签研究相关的复杂网络、标签推荐、本体、个性化、标签质量、标签聚类、标签检索等是我国图情领域社会化标签的研究热点。

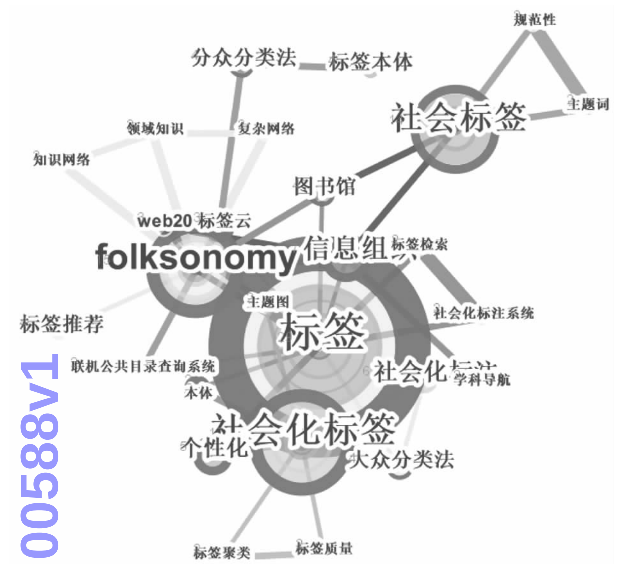


图 2　关键词共现网络示意

2.4　研究框架

通过阅读文献和社会化标签研究热点分布情况可知，标签云、主题图、图书馆、学科导航、社会化标注系统等关键词体现了社会化标签的实际应用；复杂网络、知识网络、信息组织、标签本体、主题词、领域知识是标签信息研究的内容；标签质量、规范性是标签质量研究的话题点；由于样本文献主要针对社会化标注系统进行研究，因此将个性化、标签推荐归纳入信息推荐的相关内容。综合对社会化标签的相关研究，笔者将从标签应用、标签信息、标签质量、信息推荐、信息检索 5 个研究角度对目前我国图书情报与数字图书馆学科领域的社会化标签研究情况进行探讨，以求更进一步了解社会化标签在当前学科背景下的发展动向。

3　社会化标签的应用研究

我国图情领域对社会化标签的应用研究主要集中在 2012 至 2014 年间，通过文献分析，笔者将从标签应用的领域、模式、管理对社会化标签的应用研究情况进行概述。

3.1　标签应用的领域

随着 Web2.0 信息技术的发展，社会化标签在教育、文化、科研、技术等领域都得到了广泛的应用和发展。在教育领域，图书馆是主要的标签应用场景，最早将标签引入图书馆学领域是在 2007 年，典型的应用系统如宾夕法尼亚大学图书馆的社会标签系统 PennTags、密歇根大学的 MTagger、中山大学图书馆系统。我国图情领域关于社会化标签的应用研究，主要是对高校图书馆的标签应用研究。薛娟[3]将 Tag 技术应用于学科建设服务中，设计并实现了基于社会化标签的高校图书馆学科信息导航系统，优化了学科资源的共建与共享体系；刘向红[4]将标签引入非物质遗产资源展示平台，结合大众标注与非遗的传统列表方式，构建社会化非遗数据库，以尽量多层面地展示文化资源，增强用户体验，拓展了社会化标签在文化领域的应用形式。

社会化标签不仅仅是科研领域的研究热点，其可视化形式——标签云、主题图等在学术研究、网络平台等领域也得到了广泛应用。沈奎林等[5]综合读者的文献需求及馆藏文献特点，以标签云的形式展示研究文献的关键词，提供了更人性化的图书检索和信息服务方式。也有学者通过提取社交网络数据构建标签云图，进行相关功能的拓展或优化。例如，毕强等[6]选取豆瓣网中图情领域图书标签数据绘制标签云图，展示标签资源的聚合结果，结合社会网络分析方法挖掘标签语义及层级关系，拓展了社会化标注系统的资源聚合维度；陈婷等[7]结合社会化标签和主题图实现了领域知识有序结构的可视化，构建了结构化的知识语义网络及语义标签资源导航，揭示了标签主题词间的关联，弥补了标签云在信息展示层面的缺陷。

3.2　标签应用的模式

标签应用模式可分为结合式、内嵌式、混合式、延伸式 4 种应用模式。冯倩然等[8]基于图书馆图书标签应用场景对 4 种应用模式进行介绍。结合式应用模式特指将图书书目与社会化标签相结合，构建社会书目系统；内嵌式应用模式指的是将标签引入到图书馆服务中，发挥社会化标签的作用，增加图书资源建设中的用户参与度；混合式应用模式即结合多种标签应用模式，丰富图书书目系统；延伸式应用模式指的是借助成熟的社会化标签网站，扩大图书馆服务广度及深度。

标签可以有效揭示数字学术资源内容，当前大学图书馆中 Folksonomy 的应用集中在联机搜索和图书分类领域。高校图书馆引入标签主要采取嵌入式服务模式，即在图书馆集成管理系统中嵌入标签，利用标签进行用户自主标注和检索，以标签链接方式定位资源等。延伸式服务模式也是应用的一种，即借助社会标签将图书馆的服务辐射出去。

李旭晖，李媛媛，马费成. 我国图情领域社会化标签研究主要问题分析[J]. 图书情报工作,2018,62(16):120-131.

3.3 标签应用的管理

学者们主要从图书馆标签应用角度研究了标签应用的管理问题。郑燃[9]构建了基于 Folksonomy 应用的图书馆信息组织模型,以实现良好的数据处理体系。林婷[10]通过对 112 所“211”工程高校进行实例调查,运用经典扎根理论,总结了我国高校图书馆 Folksonomy 的管理机制,包括运行机制、动力机制、约束机制。由于网络活动方式的多样性、用户与管理者身份的双重型、技术平台的非单一性,图书馆标签运行机制嵌入了社会网络的管理方法,呈现多样性的特征。在用户社会心理驱动、管理主体的拉动性与推动性的作用下,图书馆标签管理的动力机制呈现出标签拉动性利益未受重视,推动性利益依赖于系统商的情况。而对于图书资源标签的约束则主要来自用户自身,在权利约束方面并不明显。

社会化标签揭示了信息资源的内容或关联信息,为信息组织、检索提供了新的应用方式。但标签在 OPAC 上的应用尚处于摸索和完善阶段,其自身的不受控、无结构、语句问题等,使它并未广泛应用于图书馆。如何更好地匹配不同学科或群体,需要图书馆及其他相关学术研究领域构建更完善的信息管理模式。

4 社会化标签的信息研究

研究者们主要从复杂网络、本体、聚类等角度对标签信息进行研究,以探讨标签间的关系与深层语义信息,为各种信息服务提供规范化的数据基础。

4.1 标签信息研究

由于用户标注的随意性,社会化标签具有一词多义等特性,使得标签的搜索准确率降低。研究者们利用知识组织工具、统计方法或数学模型,引入标签消歧机制,以挖掘标签的真实信息。窦玉萌[11]将标签消歧方法分为基于数据挖掘方法、基于统计分析方法、利用相关知识组织工具、引入控制机制和开发可视化组件消歧 5 类,并从用户参与度、消歧时机、消歧性质、实验与应用情况、发展前景方面进行比较分析,指出可视化组件消歧方法具有较好的发展前景。

标签本体体现了标签、用户、资源之间的关系,并被广泛应用于标签信息表示研究上。祝贺[12]通过分析标签本体多元标注形式,在研究本体构建方法基础上,提出了骨架法标签本体构建方法,为标签本体的应用扩展做出了理论研究。窦永香等[1]通过对本体和社会化标签文献的关联研究情况的可视化分析,指出本体与社会化标签关联多个学科领域,包括知识共享、语义网、用户生成内容、标签本体、标签推荐等。白华[13]采用标签——概念映射和集成技术把标签映射到概念框架中,从而构建兼有大众分类和概念本体特征的新本体,在本体容纳性、灵活性、多向语义关联方面都显著优于传统分类线性结构。大众分类对知识组织系统和其他本体的影响主要体现在三个方面,包括资源决定体系、用户因素、分布式资源和用户标注造成的语义难题,所以基于语义规则聚合内容成为关联本体中不同标签的关键手段。白华[14]结合大众分类本体和一般概念本体,构造一个以概念为中心的集成本体,成为整合个人标签和知识框架及语义信息组织中较好的方法。

社会化标签系统中的标签共现关系揭示了标签的语义特征。Folksonomy 中用户标签个体间的语义紧密程度接近于随机网络的自由松散化分布特征,整体的语义关系方面拥有趋近于领域本体的较高的语义紧密程度[15]。李慧宗等[16]提出基于综合共现相似度的标签共现谱聚类方法,通过基于标签共现关系的标签聚类,精确刻画出标签语义的相似性,有效揭示了标签的潜在主题、概念及语义结构。

4.2 标签间关系研究

标签所具有的丰富语义使得标签间具有多种关系,如标签共现关系、同义词关系、上下术语关系、相关关系等,加强对标签间的属性、关系的挖掘有利于完善标签资源聚合体系,实现标签资源的优化重组。基于本体的标签间语义关系的研究,弥补了社会化标签在语义准确度、信息智能处理的不足。丁婉莹[17]综合用户标签在构建受控词表和提供个性化服务方面的优势,提出了一种基于用户标签的个人本体构建的流程,规范了用户标签的词汇、词间关系;陈开慧[18]构造了本体和分众分类融合模型,以提高标签资源检索性能;熊回香、王学东[19]通过探讨标签概念空间与领域本体间的映射机制,实现了对标签的语义组织,优化了社会化标签导航及搜索机制。

标签共现是在大规模信息资源的集合中,通过该语义关联和内容关联来挖掘有相关联系的标签[20]。研究者结合标签共现与社会网络分析方法对标签关系进行分析。李亚婷等[21]通过对 CiteUlike 数据进行实证研究,构建了以标签为单位的共现网络,分析认为标签共现网络呈现“小世界现象”与“无标度特征”,并提出构建资源的关联网络,以此优化基于标签的信息资源推荐;商宪丽等[20]通过对学术博客标签资源进行社会网络分析,指出学术博客博文资源之间的社会网络

性质,包括“小世界特性”和“无标度网络”性质。

社会网络分析法(SNA)通过对标签关联关系的挖掘和群组划分,表现出极强的逻辑性、层次性和整体性。在样本文献中,Y. I. Ming 等[22]最早提出将社会网络分析方法应用于标签分析中,并提出 4 种类型的分析方法,包括点－线－密度分析方法、中心性分析方法、凝聚子群分析方法、结构化方程分析方法;吴江[23]根据自由分类标签和资源的关系连接,提出凝聚子群分析构建自动分类网络地图,以解决标签等级结构不明确的问题;孙中秋等[24]采集豆瓣图书情报学科图书标签数据,通过对比研究基于 Folksonomy 和基于 SNA 的社会化标注系统标签资源聚合效果,指出 SNA 在实现标签资源聚合的维度、技术、效果、应用上都更具优势;易明等[25]应用网络分析的基本原理,构建了社会化标签系统中基于社会网络的知识推送网络演化总体模型,拓展了社会网络的研究领域;滕广青等[26]采用复杂网络分析方法构建了社群知识网络中的知识群落,并分析出 Folksonomy 模式中的知识群落具有动态发展、结构嵌套、知识关联丰富等特性;李锦霞[27]通过构建共标签和关联标签网络体系,结合社会网络分析方法得出数字图书馆馆藏资源的主题特征,为数字图书馆资源的管理提供参考;刘雅姝等[28]基于领域知识网络路径长度划分标签局部层级,从时间序列角度对标签局部层级结构的演化进程进行了分析,指出标签的局部层级结构相较于传统知识组织体系的内部关系更为复杂多样,且随领域知识的发展不断变迁,期间伴随标签晋升、降级的情况。

4.3 标签信息行为研究

普通用户与专家的标注行为存在部分的一致性,但由于专业背景、兴趣爱好、认知特征的不同,不同用户的标注行为所引起的标签间关系的差异也引起了研究者们的关注,部分研究者从用户标注行为角度对标签关系信息进行了研究。潘婵等[29]通过分析不同领域标签和关键词的相似度,指出相同用户对资源进行标引时,其标签和关键词的相似度相差不大,而受“最省力法则”“布拉德福定律”和“老化定律”的影响,不同领域的标签和关键词相似度则有差距;李枫林等[30]分析了不同标注行为所形成的标签与被标注资源间的关系,并将相关性分析用于计算标签相关性计算上,有效提高检索结果重排的效果。

用户即是信息组织者也是信息搜寻者。姜婷婷等[31]基于豆瓣网从信息搜寻角度对用户行为进行了研究,指出在社会化标注系统中,用户对于信息搜索一般采取多种模式,具有“多策略”倾向,信息偶遇者多为访问历史较短的新用户,关联性浏览者查看的资源数量明显高于搜索者。吴丹等[32]通过对武汉大学图书馆用户图书标注行为的调查,指出社会化标注行为的个性化特征,用户标注多源于对书籍的归类及推荐需求,更倾向于使用标签而非添加标签;由于社会化标注在图书馆应用尚不普遍,用户鲜少关注与图书标签相关的讲座或公共活动。

除了对文本资源的检索,图像检索是用户获取信息资源的方式之一,图像所具有的描述属性和特征影响着用户标注行为。邓三鸿等[33]基于社会标签分类模型调查分析了华特斯艺术博物馆数字门户网站,并指出用户更偏向于运用通用词语描述图像主题内容,更倾向于描述人或物的相关内容。

5 社会化标签的质量研究

对于社会化标签的质量研究,样本文献中包括针对标签质量影响因素的分析、标签质量评估算法或模型的研究及优化,其他主要是对标签质量的案例调查分析。

5.1 标签质量的影响因素

低质量标签干扰了社会化标注系统中资源组织的秩序,降低了相关应用的有效性及用户满意度。常见的低质量标签包括:标签过于主观、拼写错误、垃圾标注等[34]。李蕾、章成志[35]将社会化标注问题分为两类:用户标注行为导致的质量问题和标注系统缺乏语义规范导致的质量问题。标注行为问题主要表现在用户标注过于主观,拼写错误、垃圾标签现象;语义规范问题表现在多义词标注、同义词标注问题。由于无法准确描述资源属性或主题,低质量标签成为干扰标签应用的重要影响因素。刘礼锋、潘旭伟[36]通过构建社会化标签主题模型,研究了社会化标注中用户标签的主题鲜明性,指出大多数用户标签鲜明性程度介于极鲜明与极模糊之间,标签主题的鲜明性随用户的活跃性增加而增强。

在社会化标注系统中,用户标注行为不受限制,加之标签规模的逐渐增加,标签质量受文化背景、语言因素等方面的影响也越来越显著。在此背景下,社会化标签存在较多近义词、多义词,仅从语法层面不足以了解标签的全部特征,如语义特征等[37]。章成志等[38]指出标签质量评估在多语种量化分析研究上仍显不足;杜友桃、徐晨辰[39]以豆瓣网为研究对象,从标注者动机和标注系统层面调查分析当前社会化标签质量的影

响因素,指出用户的广泛性、标注随意性使得标注系统显露诸多缺陷。

5.2　标签质量的评估

标签质量的评估方法可分为人工评价和自动评价两种,李蕾等[40]依据不同资源或评估参数,列举出质量评估的代表性方法,包括利用人工评价,基于标签统计属性,依据规范词语,依据标签和主题词的比较,依据用户标签与专家标注结果的比较,依据标签与文本内容关键词的比较,基于信息检索的方式,基于用户、资源、标签三者关系来评估标签质量[35]。人工评估标签质量准确度较高,但耗时较长,难以大规模应用实施;基于标签自身属性的评估可用于大规模数据的质量评估,但未考虑到用户标注动机等其他因素的影响;依据规范词语进行隐含的质量评估一般通过对用户标注进行实时评估,以标签推荐、拼写检查、输入提示等方式,减少用户误输入、垃圾标签等问题;依据标签和主题词的比较进行评估是根据用户标注与主题词库的匹配度进行质量评估的,受主题词更新速度、覆盖范围的限制,此方法的有效性并不稳定;相较于主题词表,基于专家标注结果的评估适应了实际情况,可进行及时更新;依据与文本内容关键词的比较主要通过抽取文本资源的关键词,进行相似度匹配,此方法对文本关键词抽取质量要求较高;用户标签的访问率也体现了标签对于信息资源表述的准确度,基于信息检索方式便是以用户检索、访问情况进行标签质量的评估;通过分析用户间的相似度、资源的匹配度、标签的语义相似性来对标签质量进行评估的方法考虑了用户、资源、标签之间的联系,但对这三类资源的依赖性也较强。

5.3　标签质量的优化

对社会化标签质量的优化有利于过滤信任度低的噪音标签,提高标签应用质量。针对社会化标签的自由性缺陷,标签质量控制有 4 种方法,包括在标签创建时进行控制,即用等级分类法、分面分类法控制社会性标签;在浏览标签时进行控制,即运用标签云控制社会性标签;还有依靠用户社群力量,由用户群体共同改善标签信息。

机器学习是应用较多的标签质量模型优化方法,研究者主要通过控制同义词和近义词、同名异义词和缩写词,以及通过自由标注方法的用户引导,优化系统推荐以及完善标签管理功能来弥补标签质量的缺陷。现有关于标签质量的研究大多利用标签的单一属性。章成志、李蕾[34]提出综合标签的内容属性与社会化属性,构建了支持向量机标签质量评估模型,对标签质量进行综合评估;顾晓雪等[41]通过分析不同质量的机器生成标签的聚类效果差异,优化了融合标签质量的标签聚类算法。

标签质量取决于标签与资源的相关性,典型的标签相关性判断方案包括基于标签绝对频次和基于内容的相关性判断策略。绝对频次的高低与标签相关性并不绝对成正相关,而基于内容的相关性判断策略受限于自动标引策略的召回率、策略准确率、策略通用性。何琳等[37]通过建立"社会标签－主题词"的概念空间模型,提出一种"内核受控,外壳非控"的分类模型,对社会标签进行了规范控制,以提高社会化标签的质量;林鑫、周知[42]引入活跃度指数进行标签相关性的判断,提高了关系判断策略的通用性。

规范化社会化标签是优化信息组织和检索的重要措施,在确保标注质量上起着重要作用。吴丹等[43]对比分析中英文图书标注与编目,提出规范化管理图书标签资源,包括规范并指导用户标注图书,连接社会标签与主题词表,合并图书编目生成的元数据与图书标注生成的标签,建立个性化半自动标引等方法。也有学者从标签结构角度对标签质量进行优化,以减少标签语义模糊、缺乏层次性等问题。王爽等[44]提出引入用户分类标签的方法,构建结构性的 Folksonomy 体系,以体现用户对标签及相应类目之间关系的一致性。

6　社会化标签的信息推荐研究

在基于标签的信息推荐研究文献中,标签推荐、信息推荐、推荐算法、推荐系统是出现频次较多的关键词,本部分将从标签推荐的影响因素、类型、优化三个角度对标签推荐的相关内容进行总结。

6.1　信息推荐的影响因素

将用户迫切需要的标签和资源提供给用户是社会化标注系统用户服务的宗旨与核心[6]。武慧娟等[45]从用户个性化角度指出基于标签的信息推荐的主要影响因素包括用户标注行为、用户群的粒度、用户群中的资源。由于用户标注的自由性,标签存在类目结构扁平化、语义关系混乱、语义模糊、一义多词等问题,限制了标签用于资源检索和组织的能力[46],进而影响了基于标签的信息推荐服务的效果。用户自主标注的规范性问题影响了基于本体研究领域中对标签的聚类效率[19],标签的扁平化结构也使得其在揭示资源主题及信息检索方面未被充分利用。

社会化标签与资源间的动态关系是标签推荐的影响因素之一。社会化标签通过同一标签对不同资源和

同一资源对不同标签的聚合作用来不断扩充资源主题和资源间的动态联系,标签的核心、边缘层级的演化在一定程度上反映了领域知识核心热点的更新与迭代,社会化标签与资源之间的动态关系使得标签呈现一定的规律。其中,标签使用量的"幂律分布"规律影响着个性化信息推荐效果,基于社会化标签的用户新兴趣的发掘并不适用于基于内容的信息推荐。对于协同过滤标签推荐而言,处于长尾状态的用户无法获得精准推荐也是其所面临的主要问题。针对图片、视频、音乐、网页等不同类型的标注资源,信息提取检索方法的不同也影响着信息推荐的效果。

6.2 标签推荐的类型

常见的标签推荐包括基于资源内容的推荐、协同过滤的推荐、基于关联规则的推荐。基于内容的推荐主要利用信息检索方法提取文本信息特征,并将相似度较高的信息推荐给用户。协同过滤推荐通过标签预测用户或项目的邻居节点,将其他同类对象标注的资源推荐给目标对象。基于关联规则的推荐通过提取用户所获取的信息资源之间的关系对目标用户进行类似推荐,典型案例如从客户购买商品集中挖掘出啤酒与尿布的购买模式。

武慧娟等[47]通过分析国外对个性化信息推荐的研究文献,根据用户、资源、标签组织方式的不同,将基于标签的个性化信息推荐分为三类,包括基于图论、基于张量和基于主题模型。其中,基于图论的模型研究运用图论的理论方法刻画标签及其与用户、资源之间的关系,以体现标签系统中的网络结构;基于张量的模型研究主要通过构建用户、资源、标签关系的框架模型,利用理论分析方法,从整体层面对标签资源进行研究,以优化个性推荐;基于主题模型的研究则主要从标签语义角度挖掘标签信息,匹配各种机器学习算法,提供更准确、高效的推荐信息以满足用户需求。

6.3 标签推荐的优化

在基于社会化标签的信息推荐研究领域,多种推荐方法、资源分析的组合是标签推荐优化领域中综合性较强的研究手段。张亮[48]通过对标签系统中对象关系与资源内容的融合分析,提出了基于主题模型的推荐方法;钟克吟[49]将基于用户的协同过滤算法和基于关联资源的协同过滤算法相结合,运用于高校学术资源的推荐中,体现了协同过滤对象的多样性;易明等[50]针对基于资源内容、协同过滤、关联规则的三类推荐方法的缺陷提出了基于社会化标签网络的内容推荐和基于知识互动型社会网络的协作过滤推荐,解决了社会化标签–评价矩阵方法面临的"冷开始""新用户""新资源"等问题;徐文青等[51]将书籍"热门度"因素融入基于内容的推荐算法中,以改善图书个性化推荐效果。

与传统的协同过滤算法相比,标签聚类方法在个性化服务中可有效提高推荐服务质量。标签聚类研究的核心问题在于计算标签之间的相似度,以提高标签对资源描述的准确程度。王帅腾、徐宽[52]综合了聚类和本体方面的知识,利用形式概念分析方法将标签聚类,构建基于社会化标签的文学图书本体,优化了社会化标签的资源整合。部分研究者从用户角度进行聚类,来优化标签描述效果。S. Shan 等[53]提出依据用户标注行为的代表性对用户进行排名策略,以挖掘有价值的标签,优化现有的个性化信息推荐方法;苏杨等[54]利用同义词环改进原始标签用户分类,将用户按照兴趣进行分类,优化基于群组的兴趣推送服务;李中慧[55]利用标签聚类分析构建用户兴趣模型,构建出基于标签的个性化信息服务流程框架;王娜等[56]提出融合标签权值的用户模糊聚类方法,通过构建标签的模糊相似矩阵提取标签权值,并构建标注行为的模糊相似矩阵,根据阈值计算对用户进行聚类,优化了传统的模糊聚类方法。

社会化标签的语义模糊性和标签无序性限制了标签推荐的效率,词表在解决标签主题分类、标签语义定位问题上占据着重要位置。受控词表的专业化、含义明确化特点有利于规范用户标注行为,提高标签的检索效果。同时,Folksonomy 和受控词汇可以很好地互补,增加信息组织方式,提高 OPAC 系统的检索结果。姚晓锋[57]对比分析用户标签和主题分类的优缺点,提出以受控词表配置图书馆资源的方法。主题词表的语义关系与标签间关系具有互参照性,李艳、贾君枝[58]借助受控词表中已有的语义关系,制定出了标签语义关系挖掘规则,进而构建出轻型标签本体,以提升用户检索效率。较之于图像视觉特征分析,语义标签分析具有处理效率上的优势。宋灵超等[59]利用情感词表,提出图片标签情感类型的自动分类方法,优化了基于情感特征的图像检索与利用。H. Xie 等[60]提出了基于用户情感等级的信息分析框架,为个性化信息检索提供了理论依据。周知、梁宇[46]结合了词表规范性、标准化优势与规则在自动分类方面的能力,提出了结合自建词表和句法规则的标签自动分类方法,为基于标签的资源检索、推荐提供了参考。

研究者也从其他角度对标签推荐的优化进行了研

李旭晖，李媛媛，马费成. 我国图情领域社会化标签研究主要问题分析[J]. 图书情报工作,2018,62(16):120－131.

究。胡慕海等[61]提出涌现语义适用度及其在群偏好预测中的应用模式,使得基于知识的情境敏感性群偏好预测更加精准,提升了移动场景中面向群体的信息资源推送服务的质量;胡潜等[62]通过日志分析法调用豆瓣网图书标注数据,指出图书主题对用户标签个数、类型分布有显著影响,为图书分面检索和标签推荐设计提供了参考。

7 社会化标注系统的检索研究

通过阅读文献得知,我国图情领域学者对标签检索的研究主要从社会化标注系统的标签检索出发,限于文献量较少,本部分将从社会化标注系统的标签语义检索、检索功能、检索技术三个方面对标签检索研究进行简要综述。

7.1 基于标签的语义检索

语义检索是基于知识与语义的匹配,研究者多通过增强标签的语义信息来提供社会化标注系统的信息检索能力。标签语义检索研究包括扩展标签语义与涌现标签语义两个方面,其中扩展标签语义研究主要通过利用语义工具、本体技术对标签资源语义关系进行提取或分析,涌现标签语义研究领域主要利用概率论、图论、共现网络等方法挖掘标签资源的潜在语义。宣云干等[63]提出基于潜在语义分析的社会化标注系统标签语义检索模型,改进了标签－资源矩阵权重计算方法,以提高标签语义检索效率;陈朝[64]在 Folksonomy 标签深度语义与广度语义聚合流程基础上,综合运用语义描述方法和复杂网络可视化图谱揭示双维度的标签语义聚合过程,以此提出具备广泛视角的 Folksonomy 导航架构。

7.2 基于标签的检索功能

社会化标注系统是连接用户、标签、资源的网络平台,为用户提供对网络资源进行自由分类、标注的权限,是一种以用户为导向的信息组织管理系统。社会化标注系统的三大检索功能包括词语检索、浏览检索和个性化信息推荐。用户通过关键词、标签词或关键词与标签词的联合进行基于词语的检索,但受标签、关键词质量的影响,词语检索的精细化程度不够,词语检索的多样性及信息查全率、查准率均受到限制。浏览检索包括主题导航、标签导航等,以导航栏、标签云等形式展示给用户,其主要在用户浏览信息的过程中向用户推荐资源或用户群,由于浏览检索的随意性,其不具有固定的检索效果评估标准。个性化信息推荐包括信息资源的推荐、标签推荐、相似用户推荐和信息推送,对用户行为的跟踪与判断、用户兴趣的宏观把控,有利于提高个性化服务的针对性、精准性、及时性。

7.3 基于标签的检索技术

社会化标注系统的信息检索技术包括标签相关性排序技术、标签本体技术、个性化标签搜寻技术、多媒体标签检索技术、个性化标签推荐技术、标签词义消歧技术、垃圾标签的消除技术以及标签聚类技术等,徐志玮等[65]对标签相关性排序、标签本体、个性化标签搜寻、多媒体标签检索等技术进行了阐述。基于链接分析的网页排序算法、以向量空间模型和统计模型的聚类算法、倒排索引是目前主要的标签相关性技术。研究者利用本体对大众分类法进行表示或分析,以解决标签平面性、稀疏性所带来的标签信息检索低效率的问题,但由于标签本体构建成本较高,也有学者借助在线词表、叙词表等本体工具对标签进行层次化、查询扩展等研究。社会化标注系统的个性化搜寻主要体现在推荐系统上,包括标签推荐、自由推荐、相似用户推荐,标签共现、协同过滤、标签查询扩展等被应用于个性化推荐中,以满足用户个性化搜寻需求。针对图像的标签检索是社会化多媒体标签检索技术研究较多的话题,通过改善传统的基于文本和基于内容的图像检索方式,研究者们综合图像可视化和标签共现算法对图像进行自动标注,利用标签相关性对图像进行排序,运用本体、在线词表、地理位置等对图片标签检索进行研究。

也有学者将图数据库技术应用于标签检索系统中。王慧孜等[66]将图数据库应用于标签系统的数据管理与利用上,构建了标签系统的图数据模型,通过采集 Flicker API 图片数据构建标签系统,使用图数据库语法即可构造多维度的检索分析并获取可视化网络图。在处理标签资源上,基于图数据库的标签检索的多维性、灵活性明显优于关系数据库。

8 结语

通过对我国图书馆与情报学科领域的社会化标签研究期刊文献的调查,笔者对社会化标签发展情况、研究热点进行了总结,并从标签应用、标签信息、标签质量、信息推荐、信息检索角度对学者研究问题和方法进行阐述,分析了各领域的研究现状或不足,为图情领域社会化标签研究或系统优化工作提供参考依据。

8.1 研究现状总结

对社会化标签的研究综述主要集中在 2012 至 2015 年间,学者们主要从标签应用、研究话题、文献分布等角度进行总结分析。样本文献中最早对社会标签

进行综述分析的是翟爽、宋文所著的《社会标签进展研究概述》,该研究将社会标签作为一种新的信息组织方式,概述了社会标签的应用进展及与知识组织体系的融合研究。张丛昱、张云中[67]对 CNKI 数据库中 2006 - 2015 年间国内 Folksonomy 期刊文献进行共词分析,通过构建相异矩阵、聚类分析和多维尺度分析,综述了信息组织、知识发现、标注系统、标签应用、本体、标签推荐等方面的研究热点。近期较为全面的综述研究是刘庆旭、毕强发表的《国内外 Folksonomy 研究综述》,文章对比分析了 2004 - 2015 年间国内外 Folksonomy 理论研究、技术研究、实践应用研究情况,指出国内针对分众分类理论研究的不完善,技术创新、系统应用等方面还有待丰富等问题。综合分析我国图情领域社会化标签综述研究文献可知,现有研究多从宏观社会和学术层面对社会化标签进行阐述分析,而专门针对我国图书情报学科领域研究方向的社会化标签研究综述并不多见。笔者选用依据文献量分布情况获取研究话题的方式进行综述研究框架的设计,通过对我国图情领域社会化标签研究文献的量化分析结果得知,标签应用、标签信息、标签质量、信息推荐等是我国图情领域社会化标签主要的研究内容。

研究者多采取案例调查法对社会化标签的应用现状进行研究,通过实例调查或数据对比发掘相关应用问题,并根据不同场景或理论基础提出解决方案。与标签应用研究密切相关的研究话题包括标签云、主题图、图书馆、学科导航、社会化标注系统等,研究者们多通过探讨现有标签应用体系的优化或完善方法,来提高标签服务质量。典型的社会化标签应用平台包括 Flicker 网站、Bibsonomy. org 网站、豆瓣电影、豆瓣图书、知乎话题、科学网博客、中国知网、高校图书馆资源等网络平台,这些也是标签应用研究中考察较多的研究对象。标签信息的研究从复杂网络、本体等角度进行研究,以探讨标签间的关系与深层语义信息。基于标签的信息推荐研究主要采用模型构建的方法,通过构建理论模型,加以数据验证,优化现有的信息推荐服务,也有学者从标签聚类、系统动力学模型、主题模型、本体、协同过滤算法等角度提出了信息推荐优化策略。研究者主要基于社会化标注系统对标签检索的相关内容进行分析,通过优化检索功能、检索技术,以完善现有的标签检索。与基于社会化标签的信息推荐、信息检索、标签应用、标签信息分析效果密不可分的是标签质量,标签质量的优劣对于社会化标签相关领域的研究至关重要,研究者多从标签质量评估、标签规范性等角度进行研究。

8.2 现有研究特点及不足

在对文献研究进行总结分析的过程中,发现我国图情领域对社会化标签的研究特征及问题有所显现。

在标签应用研究层面,以本体的形式或本体的相关理论来研究标签信息及标签之间的语义关系是大多数文献选取的研究方法,将本体与社会化标签的实际应用相结合的研究是此领域的空缺。基于本体理论的社会化标签研究在实际应用中的验证分析也尚显不足。此外,高校图书馆在图情领域的标签应用研究中出现频次较多,但目前标签应用水平仍然远不及网络社交平台,体现了应用研究在教育领域的应用实效性不强。作者对此情况的原因进行深入分析后发现,除了参与图情领域社会化标签标注工作的人员偏少这一表面原因之外,更深层的原因是本体方法与社会化标签之间存在"语义鸿沟"。社会化标签本质上是用户各自对事物的局部认识,而传统的本体方法要求对事物有一个一致、全面的语义描述。这种深层次的隔阂使得社会化标签方法难以在关注知识本体分类研究的传统图情专业领域中获得广泛认可和应用。

在标签推荐层面,较多学者提出了如何将新标签归纳到现有的标签体系以及如何将非稳定状态下的用户兴趣与标签体系相匹配以精准化信息推荐的问题,新标签的自动融入与用户兴趣的变化都对标签推荐的时效性和准确性要求均较高,这也是目前基于标签的信息推荐系统在满足用户个性化信息需求过程中亟待解决的。对于标签融入的问题,作者经过分析后认为,其根源在于标签所蕴含的语义信息的表示和组织方法。社会化标签是一种灵活的、非标准的语义信息表达形式,新的标签被提出时,其蕴含的语义信息需要与底层的语义信息组织的框架相融合,才能够得到有效的推荐。当前社会化标签的研究普遍关注标签推荐的现象问题,但较少深入探讨其内部语义信息组织结构和演化方法对标签推荐的影响。

在标签质量研究层面,研究者多采用豆瓣网、博客网、知乎、高校图书馆等常见的社会化标签分析对象,在一定层面上使得研究成果有一定的对比性和参考性,但单一、大众化的研究样本,加之标签质量研究模型或方法的标签属性类型限制等问题,使得社会化标签质量研究的普适性一般。社会化标签的多样性是标签质量研究的重要组成部分,而对多样化的标签进行质量分析需要建立一个一致、自洽的语义描述机制,对不同角度、不同层次的标签进行分析与度量,并以此为

基础建立普适的质量研究框架。当前这方面还缺乏深入研究。

在标签检索研究层面,样本文献中主要是对社会化标注系统平台的检索进行研究,标签主要用于精准化信息检索,而对于社会化标签信息本身的检索研究在图情领域社会化标签研究尚显不足。标签检索同样依赖于对标签蕴含的语义信息进行表示和度量。

8.3 进一步研究的建议

随着社会化标签的应用与研究的深入,本体、复杂语义的关联分析、机器学习、自然语言处理等技术的理论和方法在各个研究领域的重要性愈加显著,相关理论或方法与社会化标签研究的有机结合也体现了目前社会化标签的研究发展趋势,未来将有更多新的分析方法或理论应用于标签研究中,有利于充分挖掘社会化标签的价值,提高对用户生成信息资源的利用率。

综合对我国图情领域社会化标签研究现状的分析,笔者认为:社会化标签的进一步深入研究需要广泛引入语义信息的表示与组织方法,从而解决标签应用、推荐、质量和检索中存在的深层次问题。在标签应用研究层面,应从标签的语义表示的多样性和演化特征入手,探讨应用领域内标签变化的特点,以更好地适应不同类型对象或不同领域的标签资源管理服务。在条件支持的情况下,将合适的研究理论付诸于实践,优化当前图书馆资源服务,以用户为中心,满足用户个性化信息需求。在基于标签的信息推荐研究层面,需要结合知识表示学习等当前语义信息表示与组织的新技术,将新标签融入现有的标签体系、用户兴趣变化等问题纳入研究重点,以提高社会化标签体系的可靠性及动态效益。在标签质量研究层面,拓展研究案例分析对象的范围,引入更多的标签属性,并通过将标签在一致合理的语义框架中进行定位,以建立标签质量的度量标准,提高研究结果的普适性。在标签检索研究层面,同样需要从其语义表示和组织结构出发,研究标签的相关语义信息的内容和拓扑结构,研究标签语义的相关性度量方法,从而拓展标签检索研究领域,以充分挖掘社会化标签的信息价值。

致谢:本文对我国图情领域社会化标签问题的研究分析得到了图书情报国家级实验中心(武汉大学)支持,特此致谢。

参考文献:

[1] 窦永香,王帮金,方倩.基于Web of Science的本体和社会化标签关联研究的进展分析[J].情报杂志,2014(6):172－178.

[2] 刘光阳.CiteSpace国内应用的传播轨迹——基于2006－2015年跨库数据的统计与可视化分析[J].图书情报知识,2017(2):60－74.

[3] 薛娟.基于社会化标签的高校图书馆学科导航系统的设计与实现[J].现代图书情报技术,2010,26(11):90－93.

[4] 刘向红.大众标注在非物质文化遗产数据库中的应用[J].图书馆建设,2014(12):39－42.

[5] 沈奎林,邵波,卢明.标签云在学术资源揭示中的研究与应用——以中国知网数据库中图书情报类核心期刊论文关键词为例[J].图书馆论坛,2013,33(3):36－42.

[6] 毕强,赵夷平,孙中秋.社会化标注系统资源聚合的实证分析[J].情报资料工作,2015(5):30－37.

[7] 陈婷,胡改丽,陈福集,等.社会标注环境下的数字图书馆知识组织模型研究——基于标签主题图视角[J].情报理论与实践,2015,38(3):63－70.

[8] 冯倩然,吴丹.社会标签在图书馆中的应用模式研究[J].图书情报工作,2011,55(21):88－92.

[9] 郑燃.基于Folksonomy的图书馆信息组织应用研究[J].图书馆杂志,2011(4):19－23.

[10] 林婷.基于经典扎根理论的我国高校图书馆Folksonomy管理机制实证研究[J].图书情报工作,2015,59(16):60－67.

[11] 窦玉萌.基于网络协作标注的标签消歧方法述评[J].现代图书情报技术,2010,26(3):27－32.

[12] 祝贺.标签本体构建研究[J].图书情报导刊,2013,23(14):105－107.

[13] 白华.利用标签－概念映射方法构建多元集成知识本体研究[J].图书情报工作,2015,59(17):127－133.

[14] 白华.大众分类本体与知识组织系统融合研究[J].图书馆学研究,2016(10):53－59.

[15] 滕广青,毕达天,任晶,等.Folksonomy中用户标签的语义紧密性研究[J].现代图书情报技术,2013(12):48－54.

[16] 李慧宗,胡学钢,何伟,等.社会化标注环境下的标签共现谱聚类方法[J].图书情报工作,2014,58(23):129－135.

[17] 丁婉莹.基于用户标签的个人本体的构建模型研究[J].现代情报,2011,31(7):42－46.

[18] 陈开慧.本体与分众分类的融合模型研究[J].图书馆学研究,2013(5):73－77.

[19] 熊回香,王学东.大众分类体系中标签与本体的映射研究[J].情报科学,2014,32(3):121－126.

[20] 商宪丽,王学东,张煜轩.基于标签共现的学术博客知识资源聚合研究[J].情报科学,2016,34(5):125－129.

[21] 李亚婷,马费成.基于标签共现的社会网络分析研究[J].情报杂志,2012,31(7):103－109.

[22] MING Y I, WANG X, DENG W. Social tag network analysis and the investigation of personalized information services based on social network analysis[J]. 中国图书馆学报年刊(英文版), 2011:102－113.

[23] 吴江.自由分类标签类聚成网状分类结构研究与实现[J].图书情报知识,2011(1):75－81.

[24] 孙中秋,陈晓美,周姗姗.基于SNA的社会化标注系统标签资源聚合研究[J].图书馆学研究,2014(13):53－61,94.

[25] 易明,毛进,邓卫华,等.社会化标签系统中基于社会网络的知

识推送网络演化研究[J]. 中国图书馆学报,2014,40(2):50 - 66.
[26] 滕广青,杨明秋,田依林,等. Folksonomy 模式中的知识群落及其核心知识分析[J]. 图书情报工作,2015,59(22):124 - 129.
[27] 李锦霞. 基于关联标签的数字图书馆馆藏资源主题分类研究[J]. 图书馆学刊,2016(10): 67 - 69.
[28] 刘雅姝,常志远,姜博,等. 标签局部层级的动态演化研究[J]. 图书馆学研究,2016(13):84 - 89.
[29] 潘婵,冯利飞,丁婉莹,等. 基于标签 - 关键词的用户行为分析[J]. 情报杂志,2010, 29(3): 139 - 142.
[30] 李枫林,张景. 基于用户标注行为的相关性分析及重排序[J]. 情报理论与实践, 2010(10):57 - 61.
[31] 姜婷婷,迟宇,史敏珊. 社会性标签系统中的信息搜寻——基于豆瓣网的实证调查[J]. 图书情报工作,2013,57(21):112 - 118.
[32] DAN W U, XIAOMEI X U, WENTING Y U. Social bookmarking behaviors of college students: a survey of Wuhan University Library users[J]. 数据与情报科学学报(英文), 2013, 6(3):15 - 30.
[33] 邓三鸿,刘齐进,夏立新. 博物馆中面向图像需求表达的标签分类研究[J]. 图书情报工作,2016(2):14 - 21.
[34] 章成志,李蕾. 社会化标签质量自动评估研究[J]. 现代图书情报技术,2015,31(10):2 - 12.
[35] 李蕾,章成志. 社会化标签质量评估研究综述[J]. 现代图书情报技术,2013,29(11):22 - 29.
[36] 刘礼锋,潘旭伟. 社会化标注中用户标签的主题鲜明性研究[D]. 杭州:浙江理工大学, 2017.
[37] 何琳,万健,何娟,等. 基于社会标签的中文图书自动分类研究[J]. 现代图书情报技术, 2014,30(9):1 - 7.
[38] 章成志,赵华,李蕾,等. 中英文图片标签质量差异比较研究——以 Flickr 为例[J]. 情报理论与实践,2018(4):123 - 127.
[39] 杜友桃,徐晨辰. 网络环境下大众标注质量调查与分析[J]. 农业图书情报学刊,2014,26(6): 134 - 136.
[40] 李蕾, 章成志. 基于标签类型的社会化标签质量测评研究[C]//第二十七届全国计算机信息管理学术研讨会论文集. 南京:南京理工大学,2013:6 - 15.
[41] 章成志,顾晓雪. 区分标签质量的机器生成标签聚类研究[J]. 现代图书情报技术,2015, 31(10):22 - 29.
[42] 林鑫,周知. 基于活跃度指数的标签相关性判断研究[J]. 图书情报工作,2015,59(9):97 - 103.
[43] 吴丹,林若楠,冯倩然,等. 社会标签的规范性研究——图书标注[J]. 图书馆论坛,2012, 32(1):1 - 7.
[44] 王爽,徐行. 基于用户分类标签建立结构性的大众分类法[J]. 图书馆学研究,2011(9):73 - 76.
[45] 武慧娟,秦雯,孙鸿飞,等. 基于标签的个性化信息推荐系统动力学模型与仿真[J]. 现代情报,2016,36(3):12 - 16.
[46] 周知,梁宇. 基于词表和句法规则的社会化标签分类研究[J]. 情报资料工作,2017(5):63 - 69.
[47] 武慧娟,徐宝祥,周兰萍. 国外基于标签的个性化信息推荐模型研究进展[J]. 图书情报工作,2013,57(7):122 - 126.
[48] 张亮. 基于 LDA 主题模型的标签推荐方法研究[J]. 现代情报, 2016,36(2):53 - 56.
[49] 钟克吟. 基于标签与协同过滤算法的学术资源推荐系统的构建[J]. 图书馆理论与实践, 2014(9):80 - 82.
[50] 易明,邓卫华,徐佳. 社会化标签系统中基于组合策略的个性化知识推荐研究[J]. 情报科学, 2011(7):1093 - 1097.
[51] 徐文青,双林平. 融合热门度因子基于标签的个性化图书推荐算法[J]. 图书情报研究, 2015(3):82 - 86.
[52] 王帅腾,徐宽. 基于社会化标签的文学图书本体构建研究[J]. 情报探索, 2017(9):7 - 13.
[53] SHAN S, ZHANG F, WU X, et al. Ranking tags and users for content-based item recommendation using folksonomy[M]// Computing and Intelligent Systems. Berlin: Springer Berlin Heidelberg, 2011:32 - 41.
[54] 苏杨,石豪,赖雯,等. 利用同义词环改进基于 folksonomy 的用户分类[J]. 图书情报工作,2011,55(8):58 - 61.
[55] 李中慧. 基于标签的个性化信息服务研究[J]. 无线互联科技, 2015(5):116 - 117.
[56] 王娜,葛毓彬. 融合标签权值的用户模糊聚类方法研究[J]. 情报理论与实践,2016, 39(3):140 - 144.
[57] 姚晓锋. 用户标签和图书馆主题分类的比较研究——以 Librarything 为例[J]. 图书馆学研究,2010(8):63 - 66.
[58] 李艳,贾君枝. 轻型标签本体与受控词表的结合研究[J]. 数字图书馆论坛,2014(8):14 - 20.
[59] 宋灵超,黄崑. 基于社会标签的图像情感分类标注研究[J]. 图书情报工作,2016,60(21): 103 - 112.
[60] XIE H, LI X, WANG T, et al. Incorporating sentiment into tag-based user profiles and resource profiles for personalized search in folksonomy[J]. Information processing & management, 2016, 52(1):61 - 72.
[61] 胡慕海,陈艳霞,谢静,等. 社群涌现语义适用性视角的情境敏感型群偏好预测研究[J]. 情报理论与实践,2018(5):85 - 90,119.
[62] 胡潜,石宇. 图书主题对用户标签使用行为影响研究[J]. 图书情报工作,2016,60(8):106 - 112.
[63] 宣云干,朱庆华. 基于 LSA 的社会化标注系统标签语义检索研究[J]. 图书情报工作,2011, 55(4): 11 - 14,57.
[64] 陈朝. 面向双维度标签语义聚合在 Folksonomy 导航架构研究[J]. 图书馆学研究, 2015(13):68 - 73.
[65] 徐志玮,郑建瑜,晏武. 社会化标注系统的信息检索关键技术研究综述[J]. 情报杂志,2014, 33(4):148 - 152,155.
[66] 王慧孜,范炜. 图数据库在标签系统中的应用研究[J]. 数字图书馆论坛,2015(4):21 - 27.
[67] 张丛昱,张云中. 国内近十年 Folksonomy 领域研究热点与趋势[J]. 新世纪图书馆, 2016(7):91 - 96.

作者贡献说明:

李旭晖:确定研究思路及框架,修改论文;
李媛媛:收集文献资料,撰写和修改论文;
马费成:提出参考意见,修改论文。

李旭晖，李媛媛，马费成. 我国图情领域社会化标签研究主要问题分析[J]. 图书情报工作,2018,62(16):120-131.

Main Problems Analysis on the Research of Social Tags in LIS

Li Xuhui[1] Li Yuanyuan[1] Ma Feicheng[1,2]

[1] School of Information Management, Wuhan University, Wuhan 430072

[2] Center for the Studies of Information Resources of Wuhan University, Wuhan 430072

Abstract: [**Purpose/significance**] Social tag is one of the hot topic in the field of library and information science in China. In recent years of the development process, the research of social tags expands to information quality, information retrieve, information recommendation, and some other research areas in LIS. The study or discussion on the research progress of tags in LIS will serve as a theoretical underpinning for the research on social tags in future. [**Method/process**] By analyzing the papers in the field of library and information science during the year 2010 to 2017 with CiteSpace, this paper visualizes the hot topics, development process and general tendency of research on social tags, and described the theme in detail from five points of practical application, including information analysis, quality assessment, information recommendation, and information retrieve. [**Result/conclusion**] Based on the situation and problems of social tags' research, some relevant suggestions are proposed, including expanding the research scopes of social tags in the field of ontology and information retrieve, emphasizing user-centered design to meet the personalized needs, applying the theoretical fruit into practice, and so on.

Keywords: social tags library science information science